

Encluster 3.0

Media Transcode Cluster

Horacio Sanson
SkillupJapan Corp.

6th Access Nova Forum 2010



- 1 Introduction
- 2 Past: Encluster Evolution
- 3 Present: Encluster 3.0 Design and Implementation



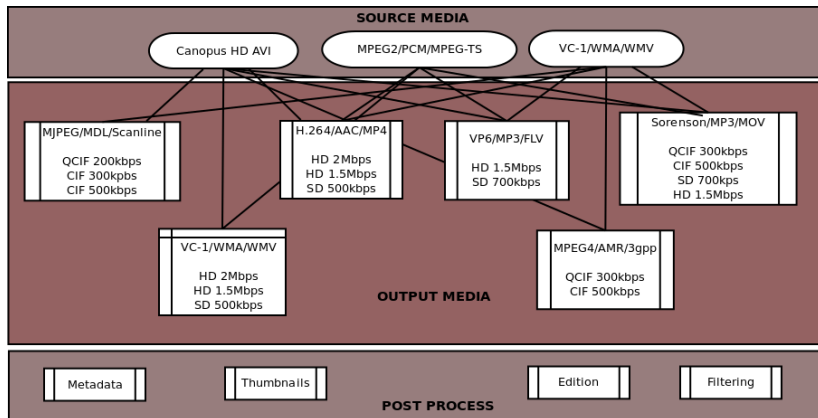
Problem Statement

Digital Distribution Problem

distribution channels \times consumer devices = Lots of media formats



Problem Statement (cont. . .)



- SkillupJapan encodes approximately 4000 media files per month.
- Media files length is between 30 minutes up to 2.5 hours.



Problem Statement (cont.)

Encoding Time per Media

About 1.5 times the media length for SD and about 2.0 times the length for HD contents.



Problem Statement (cont.)

Encoding Time per Media

About 1.5 times the media length for SD and about 2.0 times the length for HD contents.

SD Encoding Time (Lower Case)

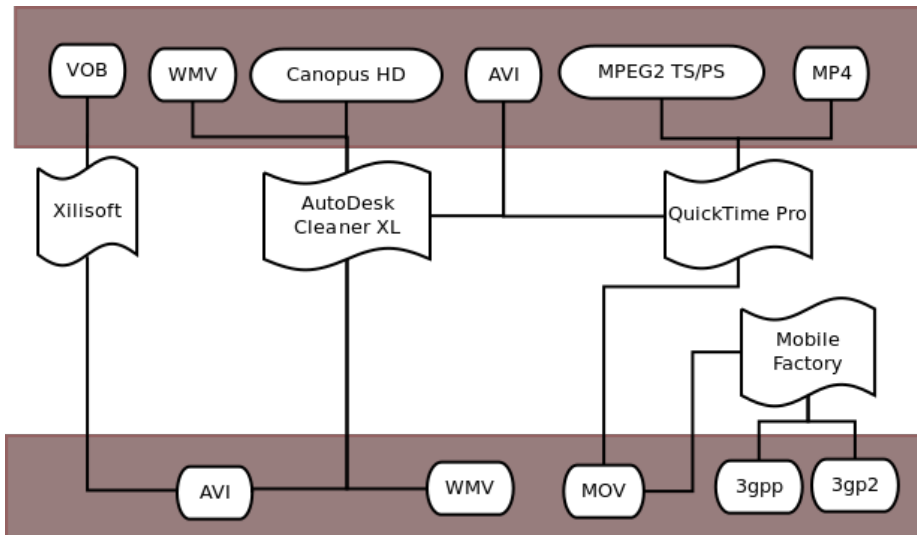
$$\frac{0.5\text{hrs} \times 1.5\text{times} \times 4000\text{movies}}{24\text{hrs/day}} = 125\text{days}$$

Required Time Resources (Worst Case)

$$\frac{2.5\text{hrs} \times 2.0\text{times} \times 4000\text{movies}}{24\text{hrs/day}} = 833.33\text{days}$$



Encluster -1.0 (Manual Encoding)



Encluster 0.1 (Scripted Encoding)

```
if [ $#= -ne 5 ]
then
echo "Usage: _encode.sh <filename> <author> <title> _..."
exit 65
fi

METADATA='-title "$3" -author "$2" -copyright "$5" -comment "$4"'
SECONDS='ffmpeg -v 3 -i $1.avi 2>&1 | awk '/Duration: [0-9][0-9]/ { spl ....''
MINUTES='ffmpeg -v 3 -i $1.avi 2>&1 | awk '/Duration: [0-9][0-9]/ { split ..''
HOURS='ffmpeg -v 3 -i $1.avi 2>&1 | awk '/Duration: [0-9][0-9]/ { split($2,..'

echo "===== _META_DATA_ ====="
echo $1 - $2 - $3 - $4 - $5
echo "$HOURS:$MINUTES:$SECONDS"

echo "===== _Creating_Thumbnail_ ====="
ffmpeg -i $1.avi -s 160x140 -an -ss 00:00:03 -an -r 1 -vframes 1 -f mjpeg \
-y $1.jpg

ffmpeg -i AKB48.avi -s 640x480 -deinterlace -pix_fmt yuv420p -g 15 -qmin 1 \
-maxrate 628000 -bufsize 628k -async 50 -acodec wmv2 -ac 2 -ab 100000 \
-vcodec wmv2 -vb 400000 -r 15 -y AKB48-640x480@500k.wmv

echo "===== _Converting_ $1_to_wmv_640x480@2Mbits_ ====="
ffmpeg -i $1.avi $METADATA -s 640x480 -deinterlace -g 15 -qmin 1 -maxrate \
1900000 -bufsize 1900k -async 50 -acodec wmv2 -ac 2 -ab 100000 -vcodec \
wmv2 -vb 1900000 -y $1-640x480@2M.wmv

echo "===== _Converting_ $1_to_wmv_640x480@1Mbits_ ====="
ffmpeg -i $1.avi $METADATA -s 640x480 -deinterlace -g 15 -qmin 1 -maxrate \
900000 -bufsize 900k -async 50 -acodec wmv2 -ac 2 -ab 100000 -vcodec \
wmv2 -vb 900000 -y $1-640x480@1M.wmv

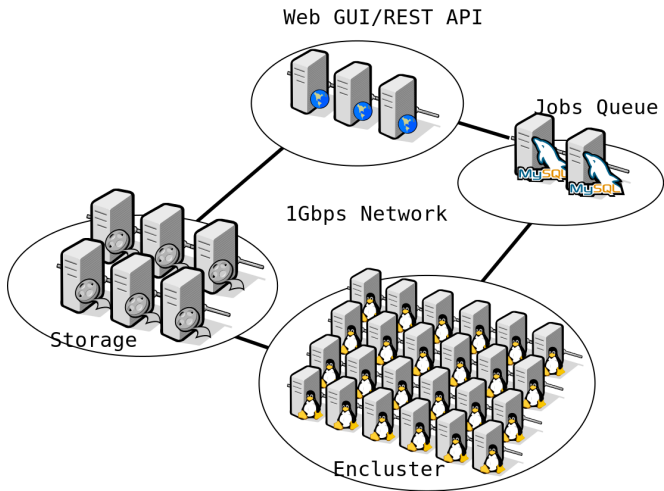
echo "===== _Converting_ $1_to_wmv_640x480@500Kbits_ ====="
ffmpeg -i $1.avi $METADATA -s 640x480 -deinterlace -g 15 -qmin 1 -maxrate \
628000 -bufsize 628k -async 50 -acodec wmv2 -ac 2 -ab 100000 -vcodec \
wmv2 -vb 400000 -y $1-640x480@500k.wmv

echo "===== _Converting_ $1_to_swf_360x270@600Kbits_ ====="
ffmpeg -pass 1 -i $1.avi -s 360x270 -acodec libmp3lame -ar 22050 -vcodec flv \
-vb 500000 -ab 96000 -f swf -y $1-360x270@600k.swf
ffmpeg -pass 2 -i $1.avi -s 360x270 -acodec libmp3lame -ar 22050 -vcodec flv \
-vb 500000 -ab 96000 -f swf -y $1-360x270@600k.swf

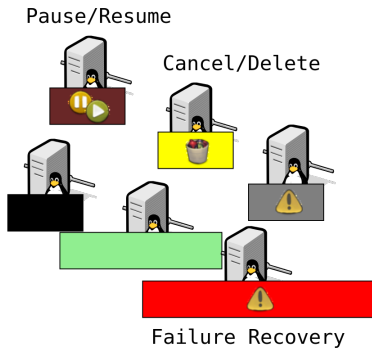
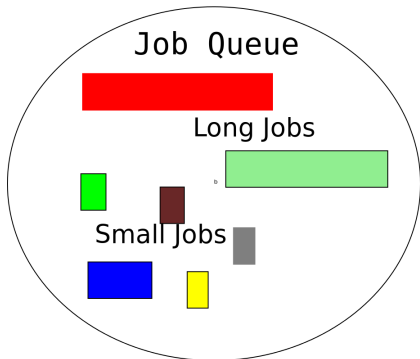
echo "===== _Converting_ $1_to_3gp_320x240@500Kbits_ ====="
```



Encluster 1.0 (Distributed Encoding)



Encluster 1.0 Limitations



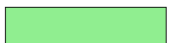
Encluster 1.0 Limitations (cont...)

Long Wait Times

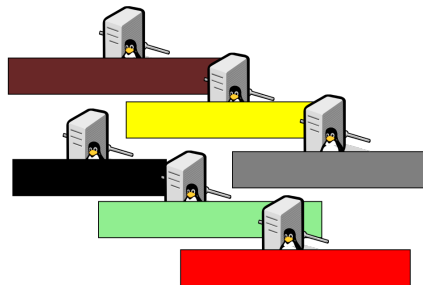
Job Queue



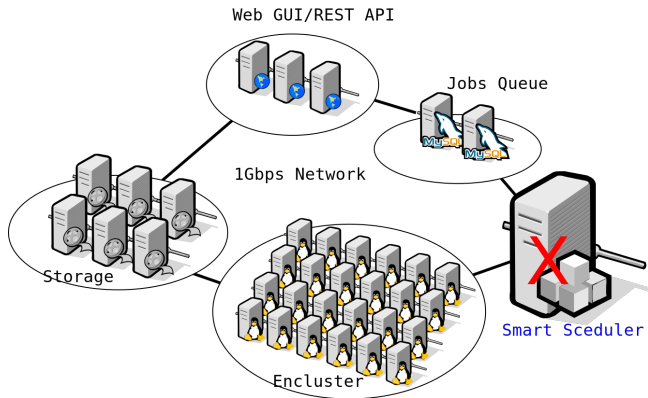
Many Long Jobs



Few Small Jobs

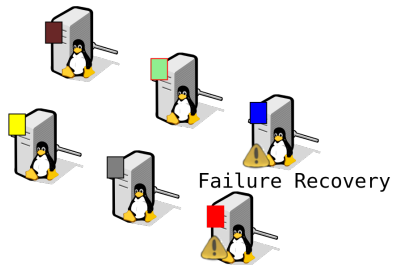


Current Solutions



Encluster 3.0: Split Encoding

Fair Task Distribution



Job Queue



Small Tasks



Cancel/Delete



Pause/Resume



Encluster 3.0: Encoding Speedup

Resolution	Length (secs)	Encoding Time (secs)	Ratio
SD (720x480)	9971.59	14957	1.5
FHD (1920x1080)	250.95	510	2.03

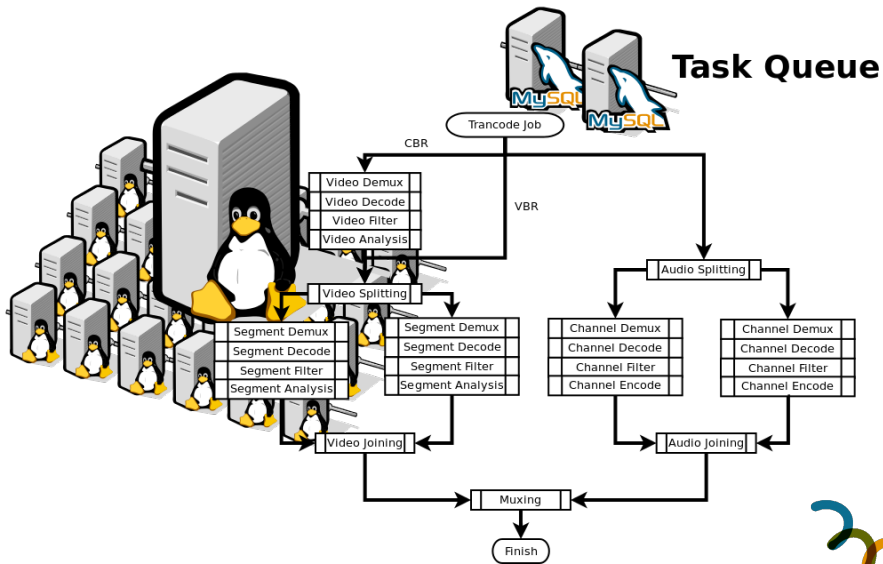
Table: Encluster 1.0 Encoding Speed

Resolution	Length (secs)	Encoding Time (secs)	Ratio
SD (720x480)	9971.59	1890	0.20
FHD (1920x1080)	250.95	231	0.92

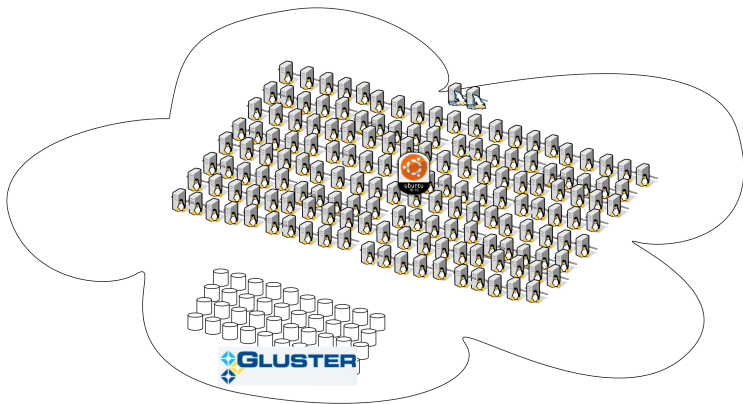
Table: Encluster 3.0 Encoding Speed



Encluster 3.0 Tasks



Encluster 3.0 is a Cloud



- High number of workers only possible in cloud systems.
- Main reason to avoid centralized schedulers.
- Implemented using distributed data storage (GlusterFS).



Future: Beyond Encluster 3.0

New Media Codecs

- Higher Resolutions: 2K, 4K, UHDT/22.2 (8K).
- New Codecs: H.265, HEVC.





Thanks For Your Attention...



Transcoding Cluster

